**Chakravarthi D** +91- 9916414037 chakri.jntukkd09@gmail.com

* Professional Data Scientist with 3.2 years of experience in Machine Learning and Business Analytics. Worked in platform development, Logistics and Telecom domains to provide ML solutions to help the business.

# STATISTICAL MODELING EXPERIENCE

* Involved in all phases of the machine learning life cycle process- research and development design, data extraction, cleaning, model building, validation and deployment.
* Developed and participated in the review of requirements, data mappings, and use cases to communicate to both business and technical teams.
* Analyzed import clearance process data and provided insights on the clearance domain.
* Developed machine learning solution on import clearance data and provides insightful resolution to reduce manual effort in the import clearance process.
* Predicted the default customers and advance problem prediction of the customer for a bank.
* Linear Regression, Logistic Regression, Decision Trees, Random Forest, Support Vector Machines and Naive Bayes Algorithm.
* Performed R&D and feasibility studies, communicate results and methodologies to the team and business stakeholders.
* Having working Knowledge on Time series, OCR and Apachi NiFi.
* Having Knowledge on tableau, H2o, Sparklyr(r package), Knime workflow and databricks tool.

# PROFESSIONAL EXPERIENCE

## *Marlabs Innovation pvt ltd*

**Senior software Engineer: Auto ML Platform development and improvements apps (Aug 2019 – Till Date)**

* Working on mAdvisor analytics platform development on Auto ML workflow.
* We tried to improve accuracy by 5% using various functions on Machine learning tasks like data preprocessing and feature engineering.
* Research and implement the new things into auto ml and testing benchmark datasets.
* Trying to added some of the algorithms into auto ml platform.
* Bug fixes if any scenarios not working on AutoML workflow.
* Creating dashboards using Tableau server.

**Commercial Invoice Extraction/Timesheet data extraction**

* Analyzed the volume of the commercial invoices by converting the pdf files into png format and doing classification by morphed templates, then extract all fields from invoice files using by master scripts.
* To extract information from invoice using by OCR Engine. Input request is exposed as Web API using by Microsoft azure api calls.
* It covers overall improvement of accuracy of their invoices and it has been automated the data extraction.
* It improved their clearance process (business), overall 4 to 6 resources efforts, process time and in terms of ROI as well.
* Trying to improve the accuracy using different models for extraction printed and handwritten text documents.
* Results comparison with google vision api calls whether it’s good or not.

***Tech Mahindra pvt ltd***

**Data Scientist: HS Code Prediction (Mar 2018 – Jul2019)**

* Predicted the most probable HS codes by using KNN and Random forest algorithms and provided insights into HS code mappings, suggested roadmap to Genius team.
* Validated the model by using K-fold cross validation and tuned parameters using GridSearchCV.
* Build end to end machine learning solution to integrate with Genius application.
* Its impact business, 60% of their HS code mapping done using this model, it reduced the 3 to 4 resources and their ROI.

**Data Analyst: Churn prediction (Telecom Client)**

* Applied business, statistical and predictive modeling expertise to identify Churn customers based on different business needs.
* Helped client to build different ML models to improve the Churn customers.
* Worked closely with client concerned team to improve the business trends.

**Machine Learning Engineer: Samsung R&D**

* Worked with Samsung R&D team to improve the Bixby functionality for calendar and social network apps.
* Performed end to end NLP (text creation, preprocessing, developing, added rules in ACE tool) task for the assistant.
* Worked with client team to improve the functionality of Bixby, partially involved to build the model and validated the model.
* Software used: Python, Spyder, Linux, MS-Excel, ACE Tool.

## *Tech Mahindra* (Nov 2016 – Feb 2018)

**Software Engineer:**

LTC Module means load traffic control , to maintain the dependencies between loads to different collections and plateforms. Accepts requests from users through a web based GUI, or directly from automated systems. It provides real time reporting of current and historic load processing. Data should be provided two ways one is manual and automated type.

* Developed Xml entries for related data.
* Responsible to developing UI Frames.
* Integration spring with database.
* Involved in bug fixing .

**HYBRIS DEVELOPER | EMBITEL TECHNOLOGIES | AUG 2015 – OCT 2016**

***Future Group Applications(BigBazar, HomeTown, eZone):***

* Developing CMS components to PLPages and Adding attributes to PLPages by using category classification.
* Involved in OMS Development Phases and Involved in Continuous Integration for Updated Data Process.
* Involved in designing the controllers in spring and Developed Xml entries for related data.
* Responsible to Integrate Spring, Dao classes and Responsible to developing facade classes for fetching data.

**Skills & Abilities**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Technical Skills** |  | R Language |  | Python |  | Artificial Intelligence |
|  |  | Data Analysis |  | Data Engineering |  | Time series, NLP |
|  |  | Machine Learning |  | Predictive Modeling |  | SQL: Intermediate |
|  |  | Data Visualization |  | Apache NiFi, Databricks |  | Java/j2ee/hybris |
|  |  |  |  |  |  |  |
| **Soft Skills** |  | Communication Skill |  | Debate |  | Presentation |

# EDUCATION

## Jawaharlal Nehru Technical University, Kakinada (Sep 2011 - Apr2014)

Master of Computer Applications, Computer Science & Engineering

## Sri Sai Baba National Degree college, SK University, Anantapur (Jun 2008-May2011)

BSc(M.E.Cs) Computer Science

# STATISTICAL MODELING PROJECTS

* *HS Code Prediction:*

To predict HSCode from shipment Description using by machine Learning Model. HSCode which is assigned by WCO organization for commodities. It is unique for each commodity. The model is trained using by historical data, AP Clustering is used to find a different cluster in the overall data and logistic regression is used to predict HSCode.

Role

Exploratory Data analysis, data cleaning and building visualizations

Count Vectorizer and TF-IDF Vectorizer used for converting text into numerical features

Built models to understand the HS codes, its usage and validated on real time data

Providing recommendations and suggestions to different shipments

Algorithms used: Random Forest, KNN, Logistic Regression, and AP Clustering

Software: Python, Spyder, RHEL, MS-Excel, MS-PPT

* *Customer ChurnPrediction*

The problem for a telecom services company is to understand why its customers are leaving. The project’s problem statement translates to predicting customer churn and understanding customer behavior based on the customer services data and demographic data.

Role

Data cleaning, understanding of feature variables, feature engineering to generate a new features and statistically identifying important features.

Built the machine learning models to predict the churn customers and validate the models

Identified Key Performance indicators (KPI's) among all the given attributes

Algorithms used: Random Forest, Logistic Regression, and Naïve Bayes

Software: R, RStudio, Windows, MS-Excel, MS-PPT